Theoretical Computer Science 123 (1994) 239–258
Elsevier

Equations and monoid varieties of dot-depth one and two*

F. Blanchet-Sadri

Department of Mathematics, University of North Carolina, Greensboro, NC 27412, USA

Communicated by M. Nivat
Received October 1991
Revised July 1992

Abstract

Blanchet-Sadri, F., Equations and monoid varieties of dot-depth one and two, Theoretical Computer Science 123 (1994) 239–258.

Each level of the Straubing's hierarchy of aperiodic monoids can be parametrized in a natural way. This paper studies this parametrization for dot-depth one and two monoids. For level one, it is shown that the mth level is defined by a *finite* sequence of equations if and only if $m=1$, 2 or 3. For level two, and for $m\geqslant 1$, a sequence of equations is given which is satisfied in the mth level and shown to ultimately define the 1st level.

1. Introduction

Let A be a given finite alphabet. The star-free languages over A are those languages which can be obtained from the finite languages over A by the boolean operations and the concatenation product. The Straubing hierarchy of star-free languages over A [18] is a hierarchy of classes of languages over A $((A^*\mathscr{V}_k)_{k\geqslant 0})$ whose union is the class of star-free languages over A. More precisely, the languages in $A^*\mathscr{V}_0$ are A^* and $\emptyset$. A language is in $A^*\mathscr{V}_{k+1}$ if it is a boolean combination of languages of the form $L_0 a_1 L_1 a_2 \dots a_n L_n$ $(n\geqslant 0)$, where the L_i's are in $A^*\mathscr{V}_k$ and the a_i's are in A. $L\subseteq A^*$ is star-free if and only if $L\in A^*\mathscr{V}_k$ for some $k\geqslant 0$ (the *dot-depth* of L is the smallest such k). The above hierarchy is closely related to the dot-depth hierarchy [8]. The

Correspondence to: F. Blanchet-Sadri, Department of Mathematics, University of North Carolina, Greensboro, NC 27412, USA. Email: blanchet@steffi.acc.uncg.edu.

* This material is based upon work supported by the National Science Foundation under Grant No. CCR-9101800. The Government has certain rights in this material. Many thanks to the referees for their valuable comments and suggestions.

0304-3975/94/$07.00
SSDI 0304-3975(92)00064-T

Straubing hierarchy of star-free languages over A is infinite for $|A|\geqslant 2$ [7, 21, 22]. Let $\mathscr{V}_k=\bigcup_A A^*\mathscr{V}_k$ and $\mathscr{V}=\bigcup_{k\geqslant 0}\mathscr{V}_k$. $\mathscr{V}$ and $\mathscr{V}_k$ are $*$-varieties of languages. Since $\mathscr{V}$ and $\mathscr{V}_k$ are $*$-varieties of languages, by a theorem of Eilenberg, there exists a monoid variety V corresponding to $\mathscr{V}$ and a monoid variety V_k corresponding to $\mathscr{V}_k$. $L\in\mathscr{V}$ ($L\in\mathscr{V}_k$) if and only if the syntactic monoid of L ($M(L)$) is in V (V_k). V is the variety of aperiodic monoids [16].

A long-standing open problem is to find out whether V_k is decidable or whether V_k can be effectively characterized. A positive answer is known only for $k=1$ [17], and partial results have been obtained for $k=2$ [19, 23].

Within each level $A^*\mathscr{V}_k$ of the Straubing hierarchy, we can consider an increasing sequence $(A^*\mathscr{V}_{k,m})_{m\geqslant 1}$ of subclasses: a language over A is in $A^*\mathscr{V}_{k,m}$ if it is a boolean combination of languages of the form $L_0a_1L_1a_2\ldots a_nL_n$ ($n\geqslant 0$) with the L_i's in $A^*\mathscr{V}_{k-1}$, the a_i's in A and $n\leqslant m$. We have $A^*\mathscr{V}_k=\bigcup_{m\geqslant 1}A^*\mathscr{V}_{k,m}$. Let $\mathscr{V}_{k,m}=\bigcup_A A^*\mathscr{V}_{k,m}$. According to Eilenberg, there exists a variety of monoids $V_{k,m}$ corresponding to the $*$-variety of languages $\mathscr{V}_{k,m}$. We have that $L\in\mathscr{V}_{k,m}$ if and only if $M(L)\in V_{k,m}$.

Eilenberg showed that every variety of monoids is *ultimately defined* by a sequence of equations and that every variety of monoids generated by a single monoid is *defined* by a (finite or infinite) sequence of equations. For example, the variety V of aperiodic monoids is ultimately defined by the equations $x^n=x^{n+1}$ ($n\geqslant 1$). The variety V_1 is ultimately defined by the equations $x^n=x^{n+1}$ and $(xy)^n=(yx)^n$ ($n\geqslant 1$). This gives a decision procedure for V_1, i.e. $M\in V_1$ if and only if for all $x,y\in M$, $x^m=x^{m+1}$ and $(xy)^m=(yx)^m$ with m the cardinality of M. The variety $V_{1,1}$ is defined by $x=x^2$ and $xy=yx$, $V_{1,2}$ by $xyzx=xyxzx$ and $(xy)^2=(yx)^2$ (the equations for $V_{1,1}$ are folklore, and the equations for $V_{1,2}$ are due to Simon), and $V_{1,3}$ by $xzyxvxwy=xzxyxvxwy$, $ywxvxyzx=ywxvxyxzx$ and $(xy)^3=(yx)^3$ [2]. Knast [11, 12] provides a sequence of equations for level one of Cohen and Brzozowski's dot-depth hierarchy.

The object of this paper is to study equations and the $V_{k,m}$'s for $k=1$ or 2. More precisely, we study the following questions.

- Can we find explicitly a *finite* sequence of equations that defines $V_{1,m}$?
- Can we find explicitly a sequence of equations that ultimately defines $V_{2,m}$?

An attempt to answer the above questions was made in [5]. There, we gave explicitly an infinite sequence of equations that defines $V_{1,m}$. We showed that for $m=1$, 2 or 3, the infinite defining sequence of equations for $V_{1,m}$ is equivalent to a finite sequence. Also, a sequence of equations that ultimately defines monoids in $V_{2,1}$ generated by two letters was given. In this paper, we first simplify the infinite defining sequence of equations for $V_{1,m}$ given in [5]. Then, we show that $V_{1,m}$ is defined by a *finite* sequence of equations if and only if $m=1$, 2 or 3 answering the first question. Also, the question of finding explicitly a sequence of equations that ultimately defines $V_{2,1}$ is solved answering the second question for $m=1$. For $m>1$, a sequence of equations satisfied in (but not necessarily complete for) $V_{2,m}$ is given. Parts of the present paper are also to be published in [5]. We have indicated those parts in the text and have repeated them for sake of completeness.

Our results are obtained by a method based on Ehrenfeucht–Fraïssé games. The idea of applying these games to study the dot-depth hierarchy dates back to the work of Thomas [21] and they were used in [1–6]. First, one regards a word $\omega \in A^*$ of length $|w|$ as a word model $w = \langle \{1, \ldots, w\}, <^w, (Q_a^w)_{a \in A} \rangle$ where the universe $\{1, \ldots, |w|\}$ represents the set of positions of letters in w, $<^w$ denotes the $<$-relation in w, and Q_a^w are unary relations over $\{1, \ldots, |w|\}$ containing the positions with letter a, for each $a \in A$. For a sequence $\bar{m} = (m_1, \ldots, m_k)$ of positive integers, where $k \geqslant 0$, the game $\mathcal{G}_{\bar{m}}(w, w')$ is played between two players I and II on the word models w and w'. A play of the game consists of k moves. In the ith move, player I chooses, in w or in w', a sequence of m_i positions; then player II chooses, in the remaining word, also a sequence of m_i positions. After k moves, by concatenating the position sequences chosen from w and w', two sequences of positions $p_1 \ldots p_n$ from w and $q_1 \ldots q_n$ from w' have been formed where $n = m_1 + \cdots + m_k$. Player II has won the play if the two subwords in w and w' given by the position sequences $p_1 \ldots p_n$ and $q_1 \ldots q_n$ coincide. If there is a winning strategy for player II in the game $\mathcal{G}_{\bar{m}}(w, w')$ to win each play, we write $w \sim_{\bar{m}} w'$. The two players play the game $\mathcal{G}_{\bar{m}}(w, w')$ on a pair of words w and w'. Player I tries to demonstrate a difference between them while player II tries to keep the words looking the same. $\sim_{\bar{m}}$ naturally defines a congruence on A^*. The importance of $\sim_{\bar{m}}$ lies in the fact that V_k can be characterized in terms of the congruences $\sim_{(m_1, \ldots, m_k)}$. Thomas [20, 21] and Perrin and Pin [14] infer that $M \in V_k$ if and only if for every morphism $\varphi: A^* \to M$ there exists $\bar{m} = (m_1, \ldots, m_k)$ such that $\sim_{\bar{m}}$ refines φ, or more precisely, $M \in V_{k,m}$ if and only if for every morphism $\varphi: A^* \to M$ there exists $\bar{m} = (m, m_1, \ldots, m_{k-1})$ such that $\sim_m$ refines φ. Hence the monoids $A^*/\sim_{\bar{m}}$ form a class of monoids that generate V in the sense that every finite aperiodic monoid is a morphic image of a monoid of the form $A^*/\sim_{\bar{m}}$.

1.1. Algebraic preliminaries

For more information on the algebraic material discussed in this section, see the books by Eilenberg [9], Lallement [13] or Pin [15].

Let A be a finite set. $|A|$ denotes the *cardinality* of A or the number of elements in A. A^*, the *free monoid* generated by A, is the set of all sequences of length $\geqslant 0$ of elements of A with concatenation being the operation (such sequences are called words). The unique string of length 0, denoted by 1 and called the empty word, acts as the identity. A *language* over A is a subset of A^*. $|w|$ denotes the length of the word w, and $w\alpha$ denotes the set of letters in w. A word u is a *prefix* (*suffix*) of w if there exists a word v such that $uv = w$ ($vu = w$). A word w' is a *factor* (or *segment*) of a word w if there exist words u and v such that $w = uw'v$. A word $a_1 \ldots a_n$ (where $a_1, \ldots, a_n$ are letters) is a *subword* of w if there exist words $w_0, \ldots, w_n$ such that $w = w_0 a_1 w_1 a_2 \ldots a_n w_n$.

An equivalence $\sim$ on A^* is a *congruence* if $w \sim w'$ implies $uwv \sim uw'v$ for all $u, v, w, w' \in A^*$. A congruence $\sim$ is *aperiodic* if there exists $n \geqslant 0$ such that $w^n \sim w^{n+1}$, for all w. The $\sim$-*class* of w is $[w]_\sim = \{w' \mid w \sim w'\}$. The set of all $\sim$-classes is denoted by $A^*/\sim$ and the *index* of $\sim$ is defined as the cardinality of $A^*/\sim$. This set becomes a monoid by considering the operation $[w]_\sim[w']_\sim=[ww']_\sim$; $[1]_\sim$ acts as identity. There exists a surjective morphism $\sim: A^* \to A^*/\sim$, defined by $w\sim=[w]_\sim$. Conversely, any morphism $\varphi: A^* \to M$ induces a congruence on A^* defined by $w\varphi w'$ if and only if $w\varphi=w'\varphi$. Note that we use the same symbol to denote the congruence and the related morphism. If φ is surjective, there exists an isomorphism between A^*/φ and M. Any A-generated monoid can then be represented as a quotient of A^* by a congruence.

If $L\subseteq A^*$ is a union of $\sim$-classes, we say that L is a *$\sim$-language*. For any language L over A, the *syntactic congruence* of L is defined by $w\sim_L w'$ if and only if for all $u, v\in A^*$, $uwv\in L$ if and only if $uw'v\in L$. $\sim_L$ is the congruence of minimal index with the property that L is a $\sim$-language, i.e., for any congruence $\sim$ on A^*, L is a $\sim$-language if and only if $\sim\,\subseteq\,\sim_L$. The quotient monoid $A^*/\sim_L$ is denoted $M(L)$ and is called the *syntactic monoid* of L. If M is a monoid and there exists a morphism $\varphi: A^*\to M$ such that $L=S\varphi^{-1}$ for some $S\subseteq M$ we say that M *recognizes* L. A language is said to be *recognizable* if it is recognized by a finite monoid. Kleene's theorem asserts that the regular languages in A^* are exactly those recognized by finite monoids. It is well known that $M(L)$ is the monoid M of minimal cardinality with the property that M recognizes L; in fact, $M(L)\prec M$ ($M(L)$ divides M or $M(L)$ is a morphic image of a submonoid of M) if and only if M recognizes L. Also L is regular if and only if $M(L)$ is finite. W is a *variety of monoids*, or M-variety, if

- it is a class of finite monoids closed under division, i.e., if $M\in W$ and $M'\prec M$, then $M'\in W$, and
- it is closed under finite direct product, i.e., if $M, M'\in W$, then $M\times M'\in W$.

For any class $\mathscr{C}$ of finite monoids, we denote by $\langle\mathscr{C}\rangle_M$ the least M-variety containing $\mathscr{C}$. Clearly, $M\in\langle\mathscr{C}\rangle_M$ if and only if there exists a finite sequence $M_1,\ldots,M_n$ of monoids of $\mathscr{C}$ such that $M\prec M_1\times\cdots\times M_n$. We call $\langle\mathscr{C}\rangle_M$ the M-variety generated by $\mathscr{C}$.

$\mathscr{W}$ is a *$*$-variety of languages* if

- for every finite alphabet A, $A^*\mathscr{W}$ denotes a class of recognizable languages over A closed under boolean operations,
- if $L\in A^*\mathscr{W}$ and $a\in A$, then $a^{-1}L=\{w\in A^* \mid aw\in L\}$ and $La^{-1}=\{w\in A^* \mid wa\in L\}$ are in $A^*\mathscr{W}$, and
- if $L\in A^*\mathscr{W}$ and $\varphi: B^*\to A^*$ is a morphism, then $L\varphi^{-1}=\{w\in B^* \mid w\varphi\in L\}$ is in $B^*\mathscr{W}$.

Eilenberg has shown that there exists a one-to-one correspondence between M-varieties and $*$-varieties. To a given $*$-variety of languages $\mathscr{W}$ corresponds the M-variety W generated by the syntactic monoids of the languages in $A^*\mathscr{W}$ for some A, and to a given M-variety W corresponds the $*$-variety of languages $\mathscr{W}$ where $A^*\mathscr{W}$ is the class of subsets L of A^* for which there is $M\in W$ such that $M(L)\prec M$. The notion of variety captures the conditions under which a family of languages can be characterized by monoids and vice versa.

Let $w, w'\in A^*$. A monoid M satisfies the equation $w=w'$ if and only if $w\varphi=w'\varphi$ for all morphisms $\varphi: A^*\to M$. One can show that the class of monoids M satisfying the equation $w=w'$ is an M-variety, denoted by $W(w, w')$. Let $(w_n, w'_n)_{n\geqslant 1}$ be a sequence of pairs of words of A^* (when considering a sequence of equations, we may need an infinite alphabet A). Consider the following M-varieties: $W' = \bigcap_{n \geqslant 1} W(w_n, w_n')$ and $W'' = \bigcup_{m \geqslant 1} \bigcap_{n \geqslant m} W(w_n, w_n')$. We say that $W'(W'')$ is *defined* (*ultimately defined*) by the equations $w_n = w_n'$ $(n \geqslant 1)$: this corresponds to the fact that a monoid M is in $W'(W'')$ if and only if M satisfies the equations $w_n = w_n'$ for all $n \geqslant 1$ (for all n sufficiently large). The equational approach to varieties is discussed in [9]. Eilenberg and Schützenberger [10] showed that every M-variety is ultimately defined by a sequence of equations and that every M-variety generated by a single monoid is defined by a sequence of equations.

2. Some combinatorial properties

The proofs of our main results in this paper rely on some combinatorial properties of the $\sim_{(m)}$'s, the $\sim_{(1,m)}$'s and the $\sim_{(n,m)}$'s stated in Sections 2.1–2.3. Parts of this section appear in [5] but are needed to understand the proofs of our new results.

2.1. *Of the $\sim_{(m)}$'s*

Simon's effective characterization of V_1 [17] depends on some combinatorial properties of the congruences $\sim_{(m)}$ stated in this section. A monoid M in V_1 satisfies $x^m = x^{m+1}$ and $(xy)^m = (yx)^m$ for some m since $M \prec A^*/\sim_{(m)}$ for some m and $x^m \sim_{(m)} x^{m+1}$ and $(xy)^m \sim_{(m)} (yx)^m$. It turns out that these two equations form a complete sequence of equations for V_1.

Lemma 2.1 ([17]). *Let $m \geqslant 1$. Let $w, w' \in A^*$.*
- *If $w \sim_{(m)} w'$, then there exists $w'' \in A^*$ such that w, w' are subwords of w'' and $w \sim_{(m)} w'' \sim_{(m)} w'$.*
- *$w \sim_{(m)} ww'$ $(w \sim_{(m)} w'w)$ if and only if there exist $w_1, \ldots, w_m \in A^*$ such that $w = w_m \ldots w_1$ $(w = w_1 \ldots w_m)$ and $w'\alpha \subseteq w_1\alpha \subseteq \cdots \subseteq w_m\alpha$.*
- *$ww' \sim_{(m)} waw'$ if and only if there exist nonnegative integers n, n', $n + n' \geqslant m$ such that $w \sim_{(n)} wa$ and $w' \sim_{(n')} aw'$ $(a \in A)$.*

2.2. *Of the $\sim_{(1,m)}$'s*

This section states some combinatorial properties of the congruences $\sim_{(1,m)}$. In this and the next sections, if $w = a_1 \ldots a_n$ is a word and $1 \leqslant p \leqslant q \leqslant n$, $w[p, q]$, $w(p, q)$, $w(p, q]$ and $w[p, q)$ will denote respectively the segments $a_p \ldots a_q$, $a_{p+1} \ldots a_{q-1}$, $a_{p+1} \ldots a_q$ and $a_p \ldots a_{q-1}$.

In the following, we talk about positions spelling the first and the last occurrences of every subword of length $\leqslant m$ of a word $w \in A^+$. Consider the following example: let $A = \{a, b, c\}$ and

$$w = \bar{a}\bar{b}\bar{b}\bar{b}b\bar{a}\bar{a}a\bar{b}\bar{c}\bar{b}\bar{c}\bar{c}c\bar{b}bb\bar{a}\bar{b}\bar{a}ba\bar{c}cbbabcca.$$

The overlined positions of w are the positions which spell the first occurrences of every subword of length $\leqslant 3$ in w.

To find the positions which spell the first occurrences of every subword of length $\leqslant m$ of a word w (or the (m) first positions in w), proceed as follows: let w_1 denote the smallest prefix of w such that $w_1\alpha = w\alpha$ (call the last position of w_1, p_1); let w_2 denote the smallest prefix of $w(p_1, |w|]$ such that $w_2\alpha = (w(p_1, |w|])\alpha$ (call the last position of w_2, p_2); ... ; let w_m denote the smallest prefix of $w(p_{m-1}, |w|]$ such that $w_m\alpha = (w(p_{m-1}, |w|])\alpha$ (call the last position of w_m, p_m). If $|w\alpha| = 1$, $p_1, \ldots, p_m$ are the (m) first positions in w and the procedure terminates. If $|w\alpha| > 1$, $p_1, \ldots, p_m$ are the (m) first position in w. To find the others, we repeat the process to find the (m) first positions in $w[1, p_1)$ and the $(m-i+1)$ first positions in $w(p_{i-1}, p_i)$ for $2 \leqslant i \leqslant m$.

A similar statement is valid to find the positions spelling the last occurrences of every subword of length $\leqslant m$ of w (or the (m) last positions in w).

The (m) first and the (m) last positions in w are called the (m) positions in w.

In the following lemmas, note that $u \sim_{(1)} v$ if and only if $u\alpha = v\alpha$.

Lemma 2.2 ([4]). *Let $m \geqslant 1$. Let $w, w' \in A^+$ and let $p_1, \ldots, p_t$ in w $(p_1 < \cdots < p_t)$ $(q_1, \ldots, q_{t'}$ in w' $(q_1 < \cdots < q_{t'}))$ be the (m) positions in w (w'). $w \sim_{(1,m)} w'$ if and only if*
- *$t = t'$,*
- *$Q_a^w p_i$ if and only if $Q_a^{w'} q_i$, $a \in A$ for $1 \leqslant i \leqslant t$, and*
- *$w(p_i, p_{i+1}) \sim_{(1)} w'(q_i q_{i+1})$ for $1 \leqslant i \leqslant t-1$.*

Lemma 2.3 ([4]). *Let $m \geqslant 1$. Let $w, w' \in A^+$ be such that $w \sim_{(1,m)} w'$. Then there exists $w'' \in A^+$ satisfying:*
- *$Q_a^w p_i$ if and only if $Q_a^{w'} p_i'$ if and only if $Q_a^{w''} p_i''$, $a \in A$ for $1 \leqslant i \leqslant t$,*
- *$w_i \sim_{(1)} w_i'' \sim_{(1)} w_i'$ for $1 \leqslant i \leqslant t-1$, and*
- *w_i, w_i' are subwords of w_i'' for $1 \leqslant i \leqslant t-1$,*

where $p_1, \ldots, p_t$ $(p_1 < \cdots < p_t)$, $p_1', \ldots, p_t'$ $(p_1' < \cdots < p_t')$, $p_1'', \ldots, p_t''$ $(p_1'' < \cdots < p_t'')$ denote the (m) positions in w, w' and w'' respectively, $w_i = w(p_i, p_{i+1})$, $w_i' = w'(p_i', p_{i+1}')$, and $w_i'' = w''(p_i'', p_{i+1}'')$ for $1 \leqslant i \leqslant t-1$.

Lemmas 2.2 and 2.3 imply that for $m \geqslant 1$, and $w, w' \in A^*$, if $w \sim_{(1,m)} w'$, then there exists $w'' \in A^*$ such that w, w' are subwords of w'' and $w \sim_{(1,m)} w'' \sim_{(1,m)} w'$.

2.3. *Of the $\sim_{(n,m)}$'s*

The following lemma gives necessary and sufficient conditions for $\sim_{(n,m)}$-equivalence.

Lemma 2.4 ([4]). *Let $m \geqslant 1$ and $n > 1$. Let $w, w' \in A^+$ and let $p_1, \ldots, p_t \in w$ $(p_1 < \cdots < p_t)$ $(q_1, \ldots, q_{t'} \in w'$ $(q_1 < \cdots < q_{t'}))$ be the (m) positions in w (w'). $w \sim_{(n,m)} w'$ if and only if*
- *$t = t'$,*
- *$Q_a^w p_i$ if and only if $Q_a^{w'} q_i$, $a \in A$ for $1 \leqslant i \leqslant t$,*

- $w(p_i, p_{i+1}) \sim_{(n-2,m)} w'(q_i, q_{i+1})$ *for* $1 \leqslant i \leqslant t-1$,
- *for* $1 \leqslant i \leqslant t-1$ *and for every* $r_1, \ldots, r_{n-1} \in w(p_i, p_{i+1})$ $(r_1 < \cdots < r_{n-1})$, *there exist* $s_1, \ldots, s_{n-1} \in w'(q_i, q_{i+1})$ $(s_1 < \cdots < s_{n-1})$ *such that*
  (1) $Q_a^w r_j$ *if and only if* $Q_a^{w'} s_j$, $a \in A$ *for* $1 \leqslant j \leqslant n-1$;
  (2) $w(r_j, r_{j+1}) \sim_{(m)} w'(s_j, s_{j+1})$ *for* $1 \leqslant j \leqslant n-2$;
  (3) $w(p_i, r_1) \sim_{(m)} w'(q_i, s_1)$;
  Also, there exist $s_1, \ldots, s_{n-1} \in w'(q_i, q_{i+1})$ *(which may be different from the positions which satisfy* (1)–(3) $(s_1 < \cdots < s_{n-1})$ *such that* 1, 2 *and*
  (4) $w(r_{n-1}, p_{i+1}) \sim_{(m)} w'(s_{n-1}, q_{i+1})$
  hold. Similarly, for every $s_1, \ldots, s_{n-1} \in w'(q_i, q_{i+1})$ $(s_1 < \cdots < s_{n-1})$, *there exist* $r_1, \ldots, r_{n-1} \in w(p_i, p_{i+1})$ $(r_1 < \cdots < r_{n-1})$ *such that* 1, 2, 3 *hold (also* 1, 2, 4 *hold)*,
- *for* $1 \leqslant i \leqslant t-1$ *and for every* $r_1, \ldots, r_n \in (p_i, p_{i+1})$ $(r_1 < \cdots < r_n)$, *there exist* $s_1, \ldots, s_n \in w'(q_i, q_{i+1})$ $(s_1 < \cdots < s_n)$ *such that*
  (5) $Q_a^w r_j$ *if and only if* $Q_a^{w'} s_j$, $a \in A$ *for* $1 \leqslant j \leqslant n$;
  (6) $w(r_j, r_{j+1}) \sim_{(m)} w'(s_j, s_{j+1})$ *for* $1 \leqslant j < n-1$;
  Similarly, for every $s_1, \ldots, s_n) \in w'(q_i, q_{i+1})$ $(s_1 < \cdots < s_n)$, *there exist* $r_1, \ldots, r_n \in w(p_i, p_{i+1})$ $(r_1 < \cdots < r_n)$ *such that* (5) *and* (6) *hold.*

3. On sequences of equations defining the $V_{1,m}$'s

Let A be an alphabet of $r+1$ letters where r denotes a nonnegative integer. In [5], for each $m \geqslant 1$, a finite sequence of equations $\mathscr{C}_{(m)}^r$ was obtained such that L belongs to $A^*\mathscr{V}_{1,m}$ if and only if its syntactic monoid $M(L)$ satisfies all the equations in $\mathscr{C}_{(m)}^r$, thus providing an effective criterion to decide if L belongs to $A^*\mathscr{V}_{1,m}$ and also implying that $\bigcup_{r \geqslant 0} \mathscr{C}_{(m)}^r$ (an infinite sequence of equations) defines $V_{1,m}$. Moreover, for $m=1, 2$ or 3, $\bigcup_{r \geqslant 0} \mathscr{C}_{(m)}^r$ was shown to be equivalent to a finite sequence of equations implying that $V_{1,1}$, $V_{1,2}$ and $V_{1,3}$ are defined by *finite* sequences of equations. Our main result in this section (Theorem 3.4) completes that result of [5] by showing that $V_{1,m}$ $(m \geqslant 4)$ is not defined by a *finite* sequence of equations. We first reduce $\mathscr{C}_{(m)}^r$ to an equivalent but simpler sequence of equations and we then show that $V_{1,m}$ is defined by a *finite* sequence of equations if and only if $m=1, 2$ or 3.

3.1. *On complete sequences of equations for the $V_{1,m}$'s*

We now state the terminology that was used in [5]. We repeat this material (that is, Definition 3.1, Theorems 3.1 and 3.2) to make the present paper self-contained. We define a new sequence of equations (Definition 3.2) and then show (Theorem 3.3) that the new sequence is equivalent to the sequence of Definition 3.1. In this paper, we work with our new Definition 3.2. Theorems 3.1–3.3 are then used to prove our main result in this section (Theorem 3.4).

By an *i-subset* we mean a subset with i elements.

Let $m \geqslant 1$. A *segment of type* $r_1(m)$ *in* $\{x_1\}$ is x_1; a *segment of type* $r_2(m)$ *in* $\{x_1, x_2\}$ is $(x_2)^e x_1$ or $(x_1)^e x_2$ for some $1 \leqslant e \leqslant m$; a *segment of type* $r_{i+1}(m)$ *in* $S_{i+1} = \{x_1, \ldots, x_i, x_{i+1}\}$ is the nonempty concatenation of at most m segments of type $r_i(m)$ in an i-subset of S_{i+1}, say S_i, followed by the concatenation (maybe empty) of at most m segments of type $r_{i-1}(m)$ in an $(i-1)$-subset of S_i, say $S_{i-1}, \ldots$, followed by the concatenation (maybe empty) of at most m segments of type $r_1(m)$ in an 1-subset of S_2, say S_1, followed by the element in $S_{i+1} - S_i$. *Segments of type* $l_i(m)$ are the mirror images, or the reversals, of the segments of type $r_i(m)$.

Definition 3.1 ([5]). Let $m \geqslant 1$ and let r be a nonnegative integer. $\mathscr{C}^r_{(m)}$ is the finite sequence of all equations of the form

$$u_r \ldots u_0 v_0 \ldots v_r = u_r \ldots u_0 x v_0 \ldots v_r,$$

where $u_0 = x^{n_0}$, $v_0 = x^{n'_0}$, where for $1 \leqslant i \leqslant r$, u_i is the concatenation of n_i segments of type $l_{i+1}(m)$ in $\{x, y_1, \ldots, y_i\}$, v_i is the concatenation of n'_i segments of type $r_{i+1}(m)$ in $\{x, z_1, \ldots, z_i\}$, and where $n_i, n'_i \geqslant 0$, $0 \leqslant i \leqslant r$, and $m = n_0 + \cdots + n_r + n'_0 + \cdots + n'_r$.

Note that $\mathscr{C}^0_{(m)}$ consists of the equation $x^m = x^{m+1}$. We have $\mathscr{C}^0_{(m)} \subseteq \mathscr{C}^1_{(m)} \subseteq \mathscr{C}^2_{(m)} \subseteq \mathscr{C}^3_{(m)} \subseteq \cdots$.

Theorem 3.1 ([5]). *Let* $|A| = r+1$, $r \geqslant 0$. *Let* M *be monoid generated by* A. *Then* M *belongs to* $V_{1,m}$ *if and only if* M *satisfies the equations in* $\mathscr{C}^r_{(m)}$. *Consequently,* $V_{1,m}$ *is defined by* $\bigcup_{r \geqslant 0} \mathscr{C}^r_{(m)}$.

Theorem 3.2 ([5]). (i) $\bigcup_{r \geqslant 0} \mathscr{C}^r_{(1)}$ *is equivalent to* $\mathscr{C}^1_{(1)}$.

(ii) $\bigcup_{r \geqslant 0} \mathscr{C}^r_{(2)}$ *is equivalent to* $\mathscr{C}^1_{(2)}$.

(iii) $\bigcup_{r \geqslant 0} \mathscr{C}^r_{(3)}$ *is equivalent to* $\mathscr{C}^2_{(3)}$.

Now, let us reduce $\mathscr{C}^r_{(m)}$ to a simpler equivalent sequence of equations $\mathscr{N}\mathscr{C}^r_{(m)}$.

Let $m \geqslant 1$. A *segment of type* $nr_1(m)$ *in* $\{x_1\}$ is x_1; a *segment of type* $nr_2(m)$ *in* $\{x_1, x_2\}$ is $(x_2)^{m_1} x_1$ or $(x_1)^{m_1} x_2$ for some m_1 where $1 \leqslant m_1 \leqslant m$; a *segment of type* $nr_{i+1}(m)$ *in* $S_{i+1} = \{x_1, \ldots, x_i, x_{i+1}\}$ is the concatenation of m_i ($m_i \geqslant 1$) segments of type $nr_i(m)$ in an i-subset of S_{i+1}, say S_i, followed by the concatenation of $m_{i-1}(m_{i-1} \geqslant 0)$ segments of type $nr_{i-1}(m)$ in an $(i-1)$-subset of S_i, say $S_{i-1}, \ldots$, followed by the concatenation of $m_1(m_1 \geqslant 0)$ segments of type $nr_1(m)$ in an 1-subset of S_2, say S_1, followed by the element in $S_{i+1} - S_i$, for some $m_1, \ldots, m_i$ where $1 \leqslant m_1 + \cdots + m_i \leqslant m$. *Segments of type* $nl_i(m)$ are the mirror images, or the reversals, of the segments of type $nr_i(m)$.

Definition 3.2. Let $m \geqslant 1$ and let r be a nonnegative integer. $\mathscr{N}\mathscr{C}^r_{(m)}$ is the subsequence of $\mathscr{C}^r_{(m)}$ of all equations of the form

$$u_r \ldots u_0 v_0 \ldots v_r = u_r \ldots u_0 x v_0 \ldots v_r$$

where $u_0 = x^{n_0}$, $v_0 = x^{n'_0}$, where u_1 is the concatenation of n_1 segments of type $nl_2(m)$ in $\{x, y_1\}$ (the last of which is of the form $x(y_1)^e$ for some $1 \leqslant e \leqslant m$), v_1 is the concatenation of n'_1 segments of type $nr_2(m)$ in $\{x, z_1\}$ (the first of which is of the form $(z_1)^e x$ for some $1 \leqslant e \leqslant m$), where for $2 \leqslant i \leqslant r$, u_i is the concatenation of n_i segments of type $nl_{i+1}(m)$ in $\{x, y_1, \ldots, y_i\}$ (the last of which starts with x if $n_1 = \cdots = n_{i-1} = 0$; otherwise, the last of which starts with a letter in $\{x, y_1, \ldots, y_j\}$ where j is the largest integer among $1, \ldots, i-1$ with $n_j \neq 0$), v_i is the concatenation of n'_i segments of type $nr_{i+1}(m)$ in $\{x, z_1, \ldots, z_i\}$ (the first of which ends with x if $n'_1 = \cdots = n'_{i-1} = 0$; otherwise, the first of which ends with a letter in $\{x, z_1, \ldots, z_j\}$ where j is the largest integer among $1, \ldots, i-1$ with $n'_j \neq 0$), and where $n_i, n'_i \geqslant 0$, $0 \leqslant i \leqslant r$, and $m = n_0 + \cdots + n_r + n'_0 + \cdots + n'_r$.

Note that $\mathcal{NC}^0_{(m)}$ consists of the equation $x^m = x^{m+1}$. We have $\mathcal{NC}^0_{(m)} \subseteq \mathcal{NC}^1_{(m)} \subseteq \mathcal{NC}^2_{(m)} \subseteq \mathcal{NC}^3_{(m)} \subseteq \cdots$.

Lemma 3.1. *Let $m \geqslant 1$ and let r be a nonnegative integer.*

- *Equations of the form*

$$w_r \ldots w_0 = w_r \ldots w_0 x,$$

$w_0 = x^{n_0}$ and, for $1 \leqslant i \leqslant r$, w_i is the concatenation of n_i segments of type $r_{i+1}(m)$ in $\{x, x_1, \ldots, x_i\}$, where $n_i \geqslant 0$, $0 \leqslant i \leqslant r$, and $m = n_0 + \cdots + n_r$ are deducible from $\mathcal{C}^r_{(m)}$.

- *Equations of the form*

$$w_0 \ldots w_r = x w_0 \ldots w_r$$

$w_0 = x^{n_0}$, and, for $1 \leqslant i \leqslant r$, w_i is the concatenation of n_i segments of type $l_{i+1}(m)$ in $\{x, x_1, \ldots, x_i\}$, and where $n_i \geqslant 0$, $0 \leqslant i \leqslant r$, and $m = n_0 + \cdots + n_r$ are deducible from $\mathcal{C}^r_{(m)}$.

Proof. We give a proof for the first claim (the second claim follows similarly). The result is trivial for $r = 0$. So assume $r \geqslant 1$. Let w'_1 be the smallest suffix of $w_r \ldots w_0$ to contain x, w'_2 the smallest suffix of $w_r \ldots w_0 - w'_1$ (here, if $w_r \ldots w_0 = uw'_1$, then $w_r \ldots w_0 - w'_1 = u$) to contain $w'_1 \alpha, \ldots$. Hence $w_r \ldots w_0 = ww'_m \ldots w'_1$, where $w \in \{x, x_1, \ldots, x_r\}^*$. There exist nonnegative integers $n'_0, \ldots, n'_r$ such that the n'_0 first segments among $w'_1, \ldots, w'_m$ are of type $l_1(m_0)$ in $\{x\}$, the n'_1 next segments are of type $l_2(m_1)$ in $\{x, y_1\}, \ldots,$ and the last n'_r segments are of type $l_{r+1}(m_r)$ in $\{x, y_1, \ldots, y_r\}$. Here, $m_0, \ldots, m_r$ are positive integers and $\{x, y_1, \ldots, y_r\} = \{x, x_1, \ldots, x_r\}$. It is possible that some of $m_0, \ldots, m_r$ be greater than m. Let $1 \leqslant k \leqslant m$. If w'_k is of type $l_{i+1}(m_i)$ $(0 \leqslant i \leqslant r)$ in $\{x, y_1, \ldots, y_i\}$ with $m_i \leqslant m$, let $w''_k = w'_k$. Otherwise, we can deduce from $\mathcal{C}^{r-1}_{(m)}$ an equation of the form $w'_k = w''_k$, where w''_k is of type $l_{i+1}(m)$ in $\{x, y_1, \ldots, y_i\}$. For example, let $r = 2$ and let

$$w'_k = y_1 x y_2 x y_2 x x x y_2 y_2 x y_2 x y_2 y_2 y_2 y_2 y_2 x y_2 y_2 x x x.$$

w'_k is of type $l_3(7)$ in $\{x, y_1, y_2\}$. Using instances of $\mathscr{C}^1_{(5)}$, one can write w'_k as w''_k where w''_k is of type $l_3(5)$ in $\{x, y_1, y_2\}$.

$$\begin{aligned}
w'_k &= y_1xy_2xy_2xxxy_2y_2[xy_2][xy_2y_2y_2y_2y_2][xy_2y_2](x)(x)x \\
&= y_1xy_2xy_2xxxy_2y_2[xy_2][xy_2y_2y_2y_2y_2][xy_2y_2](x)(x) \\
&= y_1xy_2xy_2xx[xy_2y_2][xy_2][xy_2y_2y_2y_2y_2][xy_2y_2](x)x \\
&= y_1xy_2xy_2xx[xy_2y_2][xy_2][xy_2y_2y_2y_2y_2][xy_2y_2](x) \\
&= y_1xy_2x[y_2xx][xy_2y_2][xy_2][xy_2y_2y_2y_2y_2][xy_2y_2]x \\
&= y_1xy_2x[y_2xx][xy_2y_2][xy_2][xy_2y_2y_2y_2y_2][xy_2y_2] \\
&= y_1xy_2xy_2xx[xy_2y_2][xy_2][xy_2y_2y_2y_2][y_2x](y_2)y_2 \\
&= y_1xy_2xy_2xx[xy_2y_2][xy_2][xy_2y_2y_2y_2][y_2x](y_2) \\
&= y_1xy_2x[y_2xx][xy_2y_2][xy_2][xy_2y_2y_2y_2][y_2x]y_2 \\
&= y_1xy_2x[y_2xx][xy_2y_2][xy_2][xy_2y_2y_2y_2][y_2x] \\
&= y_1x[y_2x][y_2xx][xy_2y_2][xy_2][xy_2y_2y_2y_2y_2]x \\
&= y_1xy_2xy_2xxxy_2y_2xy_2xy_2y_2y_2y_2y_2 \\
&= w''_k
\end{aligned}$$

(the segments inside [ ] are of type $l_2(5)$ in $\{x, y_2\}$ and the segments inside ( ) are of type $l_1(5)$ in $\{x\}$ or $\{y_2\}$). Let u_0 be the concatenation of the n'_0 last segments among $w''_m, \ldots, w''_1$, u_1 the concatenation of the n'_1 last segments among $w''_m, \ldots, w''_{n'_0+1}, \ldots$. So $w_r \ldots w_0 = ww'_m \ldots w'_1 = ww''_m \ldots w'_1 = wu_r \ldots u_0 = wu_r \ldots u_0x = w_r \ldots w_0x$ (here, $u_r \ldots u_0 = u_r \ldots u_0x$ is in $\mathscr{C}^r_{(m)}$) is deducible from $\mathscr{C}^r_{(m)}$. □

Theorem 3.3. *$\mathscr{C}^r_{(m)}$ is equivalent to $\mathscr{N}\mathscr{C}^r_{(m)}$. Consequently, $V_{1,m}$ is defined by $\bigcup_{r\geq 0} \mathscr{N}\mathscr{C}^r_{(m)}$.*

Proof. We show by induction on r that $\mathscr{C}^r_{(m)}$ is equivalent to $\mathscr{N}\mathscr{C}^r_{(m)}$. $\mathscr{C}^0_{(m)}$ is equivalent to $\mathscr{N}\mathscr{C}^0_{(m)}$ since both consist of the equation $x^m = x^{m+1}$. Assume that $\mathscr{C}^r_{(m)}$ is equivalent to $\mathscr{N}\mathscr{C}^r_{(m)}$. Let us show that $\mathscr{C}^{r+1}_{(m)}$ is equivalent to $\mathscr{N}\mathscr{C}^{r+1}_{(m)}$. Consider an equation in $\mathscr{C}^{r+1}_{(m)}$. It is easy to see that it has a decomposition of the form

$$wu_{r+1}u_r \ldots u_0v_0 \ldots v_rv_{r+1}w' = wu_{r+1}u_r \ldots u_0xv_0 \ldots v_rv_{r+1}w' \tag{1}$$

where $u_0 = x^{n_0}$, $v_0 = x^{n'_0}$, u_1 is the concatenation of n_1 segments of type $nl_2(n)$ in $\{x, y_1\}$ (the last segment of u_1 has x as its first letter), v_1 is the concatenation of n'_1 segments of type $nr_2(n)$ in $\{x, z_1\}$ (the first segment of v_1 has x as its last letter), and where for

$2 \leqslant i \leqslant r+1$, u_i is the concatenation of n_i segments of type $l_{i+1}(n)$ in $\{x, y_1, \ldots, y_i\}$ (the last of which starts with x if $n_1 = \cdots = n_{i-1} = 0$; otherwise, the last of which starts with a letter in $\{x, y_1, \ldots, y_j\}$ where j is the largest integer among $1, \ldots, i-1$ with $n_j \neq 0$), v_i is the concatenation of n'_i segments of type $r_{i+1}(n)$ in $\{x, z_1, \ldots, z_i\}$ (the first of which ends with x if $n'_1 = \cdots = n'_{i-1} = 0$; otherwise, the first of which ends with a letter in $\{x, z_1, \ldots, z_j\}$ where j is the largest integer among $1, \ldots, i-1$ with $n'_j \neq 0$), and where $n_i, n'_i \geqslant 0$, $0 \leqslant i \leqslant r+1$, and $m = n_0 + \cdots + n_{r+1} + n'_0 + \cdots + n'_{r+1}$, and where $w \in \{x, y_1, \ldots, y_{r+1}\}^*$ and $w' \in \{x, z_1, \ldots, z_{r+1}\}^*$ and $n \geqslant 1$.

It is possible that $n > m$. If this is the case and say, there exists a segment v of type $r_{i+1}(n)$ in $\{x, z_1, \ldots, z_i\}$ in v_i for some i, one can deduce the equation $v = v'$ where v' is of type $r_{i+1}(m)$ in $\{x, z_1, \ldots, z_i\}$ as in the proof of Lemma 3.1 (this can be done using instances of $\mathscr{C}^r_{(m)}$ or equivalently instances of $\mathscr{N}\mathscr{C}^r_{(m)}$). So we may assume that $n = m$.

We show that equation (1) can be deduced from $\mathscr{N}\mathscr{C}^{r+1}_{(m)}$. If equation (1) is already in $\mathscr{N}\mathscr{C}^{r+1}_{(m)}$, there is nothing to prove. Otherwise, there exists some segment in some u_i which is not of type $nl_{i+1}(m)$ in $\{x, y_1, \ldots, y_i\}$ or some segment in some v_i which is not of type $nr_{i+1}(m)$ in $\{x, z_1, \ldots, z_i\}$. Say some segment v in some v_i is not of type $nr_{i+1}(m)$ in $S_{i+1} = \{x, z_1, \ldots, z_i\}$. But since v is of type $r_{i+1}(m)$ in S_{i+1}, v is the concatenation of m_i $(1 \leqslant m_i \leqslant m)$ segments of type $r_i(m)$ in an i-subset of S_{i+1}, say S_i (call them $w_1, \ldots, w_{m_i}$), followed by the concatenation of m_{i-1} $(0 \leqslant m_{i-1} \leqslant m)$ segments of type $r_{i-1}(m)$ in an $(i-1)$-subset of S_i, say S_{i-1} (call them $w_{m_i+1}, \ldots, w_{m_i+m_{i-1}}$), $\ldots$, followed by the concatenation of m_1 $(0 \leqslant m_1 \leqslant m)$ segments of type $r_1(m)$ in an 1-subset of S_2, say S_1 (call them $w_{m_i+\cdots+m_2+1}, \ldots, w_{m_i+\cdots+m_1}$), followed by the element in $S_{i+1} - S_i$. If $m_1 + \cdots + m_i > m$, consider $w_1 \ldots w_m$. $w_1 \ldots w_m = w_1 \ldots w_m x$, where $x \in w_m \alpha$ can be deduced from $\mathscr{C}^r_{(m)}$ by Lemma 3.1, and hence from $\mathscr{N}\mathscr{C}^r_{(m)}$ by induction hypothesis. So $v = w_1 \ldots w_{m_1+\cdots+m_i} z = w_1 \ldots w_m z$ (where z is the element in $S_{i+1} - S_i$) is deducible from $\mathscr{N}\mathscr{C}^r_{(m)}$. If each of $w_1, \ldots, w_m$ is of type $nr_j(m)$ for some j, $v = v''$ where v'' is of type $nr_{i+1}(m)$ in $\{x, z_1, \ldots, z_i\}$ can be deduced from $\mathscr{N}\mathscr{C}^r_{(m)}$. Otherwise, repeat the process for the segments among $w_1, \ldots, w_m$ that are not of type $nr_j(m)$ (but of type $r_j(m)$) for some j. If $m_1 + \cdots + m_i \leqslant m$, $v = w_1 \ldots w_{m_1+\cdots+m_i} z$ (where z is the element in $S_{i+1} - S_i$) and there exist segments among $w_1, \ldots, w_{m_1+\cdots+m_i}$ not of type $nr_j(m)$ (but of type $r_j(m)$) for some j. Repeat the process for those segments. So $v_i = v'_i$ where v'_i is the concatenation of n'_i segments of type $nr_{i+1}(m)$ in $\{x, z_1, \ldots, z_i\}$ can be deduced from $\mathscr{N}\mathscr{C}^r_{(m)}$. A similar statement is valid for u_i. So

$$wu_{r+1} u_r \ldots u_0 v_0 \ldots v_r v_{r+1} w'$$

$$= wu'_{r+1} u'_r \ldots u'_0 v'_0 \ldots v'_r v'_{r+1} w' = wu'_{r+1} u'_r \ldots u'_0 x v'_0 \ldots v'_r v'_{r+1} w'$$

$$(u'_{r+1} u'_r \ldots u'_0 v'_0 \ldots v'_r v'_{r+1} = u'_{r+1} u'_r \ldots u'_0 x v'_0 \ldots v'_r v'_{r+1} \text{ belongs to } \mathscr{N}\mathscr{C}^{r+1}_{(m)})$$

$$= wu_{r+1} u_r \ldots u_0 x v_0 \ldots v_r v_{r+1} w',$$

so equation (1) can be deduced from $\mathscr{N}\mathscr{C}^{r+1}_{(m)}$. The result follows. $\square$

3.2. On complete finite sequences of equations for the $V_{1,m}$'s

We are now interested in the following question:

- Does there exist r' for which $\bigcup_{r\geqslant 0}\mathcal{NC}^r_{(m)}$ is equivalent to $\mathcal{NC}^{r'}_{(m)}$?

Theorem 3.2 gives a positive answer for $m=1, 2$ or 3. In this section, we give a negative answer for $m\geqslant 4$. As an application (Theorem 3.4) we show that $V_{1,m}$ is defined by a finite sequence of equations if and only if $m=$ 1, 2 or 3.

First, we give some conditions which enable us to tell if a given equation in $\mathcal{NC}^r_{(m)}$ is also in $\mathcal{NC}^{r'}_{(m)}$ with $r'<r$.

A *partition* of $\{1,\ldots,r\}$ is a collection of disjoint subsets whose union is $\{1,\ldots,r\}$ itself. An *ordered partition* is a partition in which the subsets are ordered; in this case, although the sets are ordered, the elements within the sets are not. For example, $\langle\{3,6\},\{2,5\},\{1,4,7\}\rangle$ is an ordered partition of $\{1,2,3,4,5,6,7\}$.

Let $m\geqslant 1$ and let r be a positive integer. Let

$$u_{r_1}\ldots u_0v_0\ldots v_{r_2}=u_{r_1}\ldots u_0xv_0\ldots v_{r_2} \tag{2}$$

be an equation in $\mathcal{NC}^r_{(m)}$, where $1\leqslant r_1, r_2\leqslant r$, u_{r_1} and v_{r_2} are nonempty. If there exist

- an ordered partition $\Pi_1=\langle A_1,A_2,\ldots,A_{k_1}\rangle$ of $\{1,\ldots,r_1\}$ with $A_k=\{a_{k1},\ldots,a_{kj_k}\}$ for $1\leqslant k\leqslant k_1 (j_1+j_2+\cdots+j_{k_1}=r_1)$,
- an ordered partition $\Pi_2=\langle B_1,B_2,\ldots,B_{k_2}\rangle$ of $\{1,\ldots,r_2\}$ with $B_k=\{b_{k1},\ldots,b_{kj'_k}\}$ for $1\leqslant k\leqslant k_2$ $(j'_1+j'_2+\cdots+j'_{k_2}=r_2)$,
- an ordering $a_{k_1}<\cdots<a_{kj_k}$ of the elements in A_k for $1\leqslant k\leqslant k_1$,
- an ordering $b_{k_1}<\cdots<b_{kj'_k}$ of the elements in B_k for $1\leqslant k\leqslant k_2$,
- an equation of the form

$$u'_{k_1}\ldots u'_0v'_0\ldots v'_{k_2}=u'_{k_1}\ldots u'_0xv'_0\ldots v'_{k_2} \tag{3}$$

in $\mathcal{NC}^{\max\{k_1,k_2\}}_{(m)}$ (here u'_{k_1} and v'_{k_2} are nonempty),

such that if we replace in (3) y_i by $y_{a_{i1}}\ldots y_{a_{ij_i}}$ for $1\leqslant i\leqslant k_1$, and z_i by $z_{b_{i1}}\ldots z_{bij'_i}$ for $1\leqslant i\leqslant k_2$, we get (2), then (2) is said to be *deducible by partitions* from (3).

Similarly, if $r_1=0$ or $r_2=0$.

Theorem 3.4. *$V_{1,m}$ is defined by a finite sequence of equations if and only if $m=1,2$ or 3.*

Proof. Using Theorems 3.1 and 3.2, if suffices to show that for $m\geqslant 4$, $V_{1,m}$ is not defined by a finite sequence of equations. Let $m\geqslant 4$. Assume that $V_{1,m}$ is defined by a finite sequence of equations. By Theorem 3.3, since $V_{1,m}$ is defined by $\bigcup_{r\geqslant 0}\mathcal{NC}^r_{(m)}$, we have that $\bigcup_{r\geqslant 0}\mathcal{NC}^r_{(m)}$ is equivalent to a finite sequence of equations. So there exists $r'\geqslant 0$ such that $\bigcup_{r\geqslant 0}\mathcal{NC}^r_{(m)}$ is equivalent to $\mathcal{NC}^{r'}_{(m)}$. Obviously, $r'>0$. Consider the following equation in $\mathcal{NC}^{r'+1}_{(m)}$:

$$\begin{aligned}&x^{m-2}z_1\ldots z_{r'}z_{r'+1}xz_2z_4\ldots z_{r'+1}z_1z_3\ldots z_{r'}x\\&=x^{m-1}z_1\ldots z_{r'}z_{r'+1}xz_2z_4\ldots z_{r'+1}z_1z_3\ldots z_{r'}x\end{aligned} \tag{4}$$

if r' is odd,

$$x^{m-2}z_1 \dots z_{r'}z_{r'+1}xz_2z_4 \dots z_{r'}z_1z_3 \dots z_{r'+1}x$$
$$=x^{m-1}z_1 \dots z_{r'}z_{r'+1}xz_2z_4 \dots z_{r'}z_1z_3 \dots z_{r'+1}x \tag{5}$$

if r' is even.

We show that equation (4) is not deducible from $\mathcal{N}\mathcal{C}^{r'}_{(m)}$ (the proof for equation (5) is similar). Rewrite equation (4) as $x^{m-2}w=x^{m-1}w$, where

$$w=z_1 \dots z_{r'}z_{r'+1}xz_2z_4 \dots z_{r'+1}z_1z_3 \dots z_{r'}x.$$

Let w' be the result of inserting (or deleting) a variable among $x, z_1, \dots, z_{r'+1}$ somewhere in w. We have $x^{m-2}w \not\sim_{(m)} x^{m-2}w'$ and $x^{m-1}w \not\sim_{(m)} x^{m-1}w'$. Hence no equation in $\mathcal{N}\mathcal{C}^{r'}_{(m)}$ can be used to insert (or delete) such a variable somewhere in w (here we use the fact that $m \geqslant 4$). So the only equations in $\mathcal{N}\mathcal{C}^{r'}_{(m)}$ that can be used are the ones that can insert x in front of $x^{m-2}w$ (or somewhere before w). So the only equations in $\mathcal{N}\mathcal{C}^{r'}_{(m)}$ that could be useful are: $x^{m-2}v_iv_j=x^{m-1}v_iv_j$ where v_i is a segment of type $nr_{i+1}(m)$ in $\{x, z_1, \dots, z_i\}$, and where v_j is a segment of type $nr_{j+1}(m)$ in $\{x, z_1, \dots, z_j\}$, $1 \leqslant i \leqslant j \leqslant r'$. Hence, if equation (4) is deducible from $\mathcal{N}\mathcal{C}^{r'}_{(m)}$, it is deducible by partitions from $\mathcal{N}\mathcal{C}^{r'}_{(m)}$. So rewrite $z_1 \dots z_{r'}z_{r'+1}$ as $w_1 \dots w_n$ where $n \leqslant r'$. No permutation of $w_1, \dots, w_n$ leads to $z_2z_4 \dots z_{r'+1}z_1z_3 \dots z_{r'}$. The result follows. □

4. On sequences of equations for the $V_{2,n}$'s

Our main result in this section is Theorem 4.1, which gives an equational characterization of $V_{2,1}$. This new result builds upon an equational characterization of the monoids in $V_{2,1}$ generated by two letters (which appears in [5]).

4.1. On a sequence of equations ultimately defining $V_{2,1}$

In this section, we give a sequence of equations ultimately defining $V_{2,1}$.

Definition 4.1. Let $m \geqslant 1$ and let r be a nonnegative integer. $\mathcal{C}^r_{(1,m)}$ is the finite sequence of all equations of the form

$$u_r \dots u_0xv_0 \dots v_r=u_r \dots u_0x^2v_0 \dots v_r$$

where $u_0=x^{n_0}$, $v_0=x^{n'_0}$, where for $1 \leqslant i \leqslant r$, u_i is the concatenation of n_i segments of type $l_{i+1}(2m+1)$ in $\{x, y_1, \dots, y_i\}$, v_i is the concatenation of n'_i segments of type $r_{i+1}(2m+1)$ in $\{x, z_1, \dots, z_i\}$, and where $n_i, n'_i \geqslant 0$, $0 \leqslant i \leqslant r$, and $m=n_0+\cdots+n_r=n'_0+\cdots+n'_r$.

$\mathcal{D}^r_{(1,m)}$ is the finite sequence of all equations of the form

$$u_r \dots u_0xyv_0 \dots v_r=u_r \dots u_0yxv_0 \dots v_r$$

where u_0 is the concatenation of n_0 segments of type $l_2(2m+1)$ in $\{x, y\}$, v_0 is the concatenation of n'_0 segments of type $r_2(2m+1)$ in $\{x, y\}$, where for $1 \leqslant i \leqslant r$, u_i is the concatenation of n_i segments for type $l_{i+2}(2m+1)$ in $\{x, y, y_1, \ldots, y_i\}$, v_i is the concatenation of n_i' segments of type $r_{i+2}(2m+1)$ in $\{x, y, z_1, \ldots, z_i\}$, and where $n_i, n_i' \geqslant 0$, $0 \leqslant i \leqslant r$, and $m = n_0 + \cdots + n_r = n_0' + \cdots + n_r'$.

Note that $\mathscr{C}^0_{(1,m)}$ consists of the equation $x^{2m+1} = x^{2m+2}$. We have $\mathscr{C}^0_{(1,m)} \subseteq \mathscr{C}^1_{(1,m)} \subseteq \mathscr{C}^2_{(1,m)} \subseteq \cdots$, and $\mathscr{D}^0_{(1,m)} \subseteq \mathscr{D}^1_{(1,m)} \subseteq \mathscr{D}^2_{(1,m)} \subseteq \cdots$.

Theorem 4.1. *Let $|A| = r+2$, $r \geqslant 0$. Let M be a monoid generated by A. Then M belongs to $V_{2,1}$ if and only if M satisfies the equation in $\mathscr{C}^{r+1}_{(1,n)} \cup \mathscr{D}^r_{(1,n)}$ for all n sufficiently large. Consequently, $V_{2,1}$ is ultimately defined by $\bigcup_{r \geqslant 0} (\mathscr{C}^{r+1}_{(1,n)} \cup \mathscr{D}^r_{(1,n)})_{n \geqslant 1}$.*

Proof. We have to prove that $M \in V_{2,1}$ if and only if M satisfies the equations in $\mathscr{C}^{r+1}_{(1,m)} \cup \mathscr{D}^r_{(1,n)}$ for all n sufficiently large. It is easily seen, using Lemma 2.2, that $M \in V_{2,1}$ satisfies $\mathscr{C}^{r+1}_{(1,m)} \cup \mathscr{D}^r_{(1,m)}$ for some $m \geqslant 1$. This comes from the fact that if $M \in V_{2,1}$, then M divides $A^*/\sim_{(1,m)}$ for some $m \geqslant 1$. Since $A^*/\sim_{(1,m)}$ satisfies $\mathscr{C}^{r+1}_{(1,m)} \cup \mathscr{D}^r_{(1,m)}$, M satisfies $\mathscr{C}^{r+1}_{(1,m)} \cup \mathscr{D}^r_{(1,m)}$. Moreover, M satisfies $\mathscr{C}^{r+1}_{(1,n)} \cup \mathscr{D}^r_{(1,n)}$ for all $n \geqslant m$ since $\sim_{(1,n)} \subseteq \sim_{(1,m)}$ for those n.

Conversely, let $\varphi: A^* \to M$ be a surjective morphism satisfying $w\varphi = w'\varphi$ for every equation $w = w'$ in $\bigcap_{n \geqslant m} \mathscr{C}^{r+1}_{(1,n)} \cup \mathscr{D}^r_{(1,n)}$ for some $m \geqslant 1$. Let us show that $M \in V_{2,1}$. It is sufficient to prove that for all f and g in A^*, $f \sim_{(1,m)} g$ implies $f\varphi = g\varphi$. For $f = g = 1$, it is certainly true. So assume $f, g \neq 1$ and $f \sim_{(1,m)} g$. We want to show that $f\varphi = g\varphi$. Let $p_1, \ldots, p_t$ $(p_1 < \cdots < p_t)$ $(q_1, \ldots, q_t$ $(q_1 < \cdots < q_t))$ be the (m) positions in $f(g)$. Lemma 2.3 allows us to consider only the case where

$$Q^f_a p_i \text{ if and only if } Q^g_a q_i,\ a \in A \text{ for } 1 \leqslant i \leqslant t,$$

$$f_i \sim_{(1)} g_i, \text{ and}$$

$$f_i \text{ is a subword of } g_i \text{ for } 1 \leqslant i \leqslant t-1,$$

where $f_i = f(p_i, p_{i+1})$, $g_i = g(q_i, q_{i+1})$ for $1 \leqslant i \leqslant t-1$. Here $f = a_1 f_1 a_2 f_2 \ldots a_{t-1} f_{t-1} a_t$, $g = a_1 g_1 a_2 g_2 \ldots a_{t-1} g_{t-1} a_t$, where $Q^f_{a_i} p_i$ and $Q^g_{a_i} q_i$ for some $a_i \in A$, $1 \leqslant i \leqslant t$. The above permits to consider only the case where

$$f = a_1 f_1 a_2 f_2 \ldots f_{i-1} a_i f_i a_{i+1} f_{i+1} \ldots a_{t-1} f_{t-1} a_t,$$

$$g = a_1 f_1 a_2 f_2 \ldots f_{i-1} a_i g_i a_{i+1} f_{i+1} \ldots a_{t-1} f_{t-1} a_t,$$

where f_i is a subword of g_i and $f_i \sim_{(1)} g_i$ for some i between 1 and $t-1$. We observe also that if f_i is a subword of h_i and h_i a subword of g_i, we have also $f_i \sim_{(1)} h_i$. Hence we have only to consider the case where

$$f = a_1 f_1 a_2 f_2 \ldots f_{i-1} a_i u v a_{i+1} f_{i+1} \ldots a_{t-1} f_{t-1} a_t,$$

$$g = a_1 f_1 a_2 f_2 \ldots f_{i-1} a_i u a v a_{i+1} f_{i+1} \ldots a_{t-1} f_{t-1} a_t$$

for some i between 1 and $t-1$, some a in u or in v. Assume a is in u, so $uv = u'au''v$ for some $u', u'' \in A^*$ (the proof when a is in v is similar). If $u'' = 1$, then $uv = u'av$ and

$uav = u'a^2v$. From the choice of the p_i's,

$$a_1 f_1 a_2 f_2 \dots f_{i-1} a_i u' \sim_{(m)} a_1 f_1 a_2 f_2 \dots f_{i-1} a_i u' a,$$

$$v a_{i+1} f_{i+1} \dots a_{t-1} f_{t-1} a_t \sim_{(m)} a v a_{i+1} f_{i+1} \dots a_{t-1} f_{t-1} a_t.$$

By Lemma 2.1, there exist $w_1, \dots, w_m, w'_1, \dots, w'_m \in A^*$ such that

$$a_1 f_1 a_2 f_2 \dots f_{i-1} a_i u' = w_m \dots w_1,$$

$$v a_{i+1} f_{i+1} \dots a_{t-1} f_{t-1} a_t = w'_1 \dots w'_m,$$

$\{a\} \subseteq w_1\alpha \subseteq \cdots \subseteq w_m\alpha$, $\{a\} \subseteq w'_1\alpha \subseteq \cdots \subseteq w'_m\alpha$. We can choose w_1 to be the smallest suffix of $a_1 f_1 a_2 f_2 \dots f_{i-1} a_i u'$ to contain a, w_2 the smallest suffix of $a_1 f_1 a_2 f_2 \dots f_{i-1} a_i u' - w_1$ to contain $w_1\alpha, \dots, w'_1$ the smallest prefix of $v a_{i+1} f_{i+1} \dots a_{t-1} f_{t-1} a_t$ to contain a, w'_2 the smallest prefix of $v a_{i+1} f_{i+1} \dots a_{t-1} f_{t-1} a_t - w'_1$ to contain $w'_1\alpha, \dots$ Hence $w_m \dots w_1$ is in fact a suffix of $a_1 f_1 a_2 f_2 \dots f_{i-1} a_i u'$ and $w'_1 \dots w'_m$ is in fact a prefix of

$$v a_{i+1} f_{i+1} \dots a_{t-1} f_{t-1} a_t.$$

There exist nonnegative integers $n_0, \dots, n_{r+1}$ such that the n_0 first segments among $w_1, \dots, w_m$ are of type $l_1(m_0)$ in $\{a\}$, the n_1 next segments are of type $l_2(m_1)$ in $\{a, b_1\}, \dots,$ and the last n_{r+1} segments are of type $l_{r+2}(m_{r+1})$ in $\{a, b_1, \dots, b_{r+1}\}$. Here, $m_0, \dots, m_{r+1}$ are positive integers, $a, b_1, \dots, b_{r+1}$ are in A. Similarly, there exist nonnegative integers $n'_0, \dots, n'_{r+1}$ such that the n'_0 first segments among $w'_1, \dots, w'_m$ are of type $r_1(m'_0)$ in $\{a\}$, the n'_1 next segments are of type $r_2(m'_1)$ in $\{a, c_1\}, \dots,$ and the last n'_{r+1} segments are of type $r_{r+2}(m'_{r+1})$ in $\{a, c_1, \dots, c_{r+1}\}$. Here, $m'_0, \dots, m'_{r+1}$ are positive integers and $c_1, \dots, c_{r+1}$ are in A. It is possible that some of $m_0, \dots, m_{r+1}$, $m'_0, \dots, m'_{r+1}$ be greater than $2m+1$. If this is the case for some m_i, say, $0 \leqslant i \leqslant r+1$, and w_k $(1 \leqslant k \leqslant m)$ is of type $l_{i+1}(m_i)$ in an $(i+1)$-subset of A, one can write $w_k\varphi$ as $w''_k\varphi$ where w''_k is of type $l_{i+1}(2m+1)$ in that subset of A. This can be done using instances of $\mathscr{C}^{r+1}_{(1,m)}$ and $\mathscr{D}^r_{(1,m)}$. For example, let $r = 1$ and $A = \{a, b, c\}$. Let

$$a_1 f_1 a_2 f_2 \dots f_{i-1} a_i u' = \overbrace{abc}^{w_4}\ \overbrace{bacacaaaccacaccccccaccaacacccaac}^{w_3}\ \overbrace{abbb}^{w_2}\ \overbrace{a}^{w_1},$$

and

$$v a_{i+1} f_{i+1} \dots a_{t-1} f_{t-1} a_t = \overbrace{cccca}^{w'_1}\ \overbrace{cbccba}^{w'_2}\ \overbrace{bccca}^{w'_3}\ \overbrace{abc}^{w'_4}.$$

$w_4 \dots w_1 a w'_1 \dots w'_4 \sim_{(1,4)} w_4 \dots w_1 a^2 w'_1 \dots w'_4$. Here, $m = 4$. There is $n_0 = 1$ segment in $a_1 f_1 a_2 f_2 \dots f_{i-1} a_i u'$ of type $l_1(m_0) = l_1(1)$ in $\{a\}$, i.e. w_1; there is $n_1 = 1$ segment of type $l_2(m_1) = l_2(3)$ in $\{a, b\}$, i.e. w_2; and there are $n_2 = 2$ segments of type $l_3(m_2) = l_3(11)$ in $\{a, b, c\}$, i.e. w_3 and w_4. Also, there is $n'_0 = 0$ segment in $v a_{i+1} f_{i+1} \dots a_{t-1} f_{t-1} a_t$ of type $r_1(m'_0) = r_1(1)$ in $\{a\}$; there is $n'_1 = 1$ segment of type $r_2(m'_1) = r_2(4)$ in $\{a, c\}$, i.e. w'_1; and there are $n_2' = 3$ segments of type $r_3(m_2') = r_3(2)$ in $\{a, b, c\}$, i.e. w_2', w_3' and w_4'. We have that $m_2 > 2(4) + 1$ and w_3 is of type $l_3(m_2)$ in $\{a, b, c\}$. By hypothesis, $\varphi: A^* \to M$ is a surjective morphism satisfying $w\varphi = w'\varphi$ for every equation $w = w'$ in $\mathscr{C}^2_{(1,4)} \cup \mathscr{D}^1_{(1,4)}$. In particular, since $u_1 x x^{n_0'} v_1 = u_1 x^2 x^{n_0'} v_1$ where u_1 is the concatenation of 4 segments of type $l_2(9)$ in $\{x, y_1\}$ and where v_1 is the concatenation of n_1' segments of type $r_2(9)$ in $\{x, z_1\}$ $(n_0' + n_1' = 4)$ belongs to $\mathscr{C}^2_{(1,4)}$, $u_1 x x^{n_0'} v_1 \varphi = u_1 x^2 x^{n_0'} v_1 \varphi$. Also, since $u_0 x y v_0 = u_0 y x v_0$ where u_0 (v_0) is the concatenation of 4 segments of type $l_2(9)$ $(r_2(9))$ in $\{x, y\}$ belongs to $\mathscr{D}^1_{(1,4)}$, $u_0 x y v_0 \varphi = u_0 y x v_0 \varphi$. Hence, one can write $w_3 \varphi$ as $w_3'' \varphi$ where w_3'' is of type $l_3(9)$ in $\{a, b, c\}$.

$$w_3\varphi = baca[caa][ac][ca][ca]c^2(c)(c)(c)(ac)\ caacacccaac\varphi$$

(using an instance of $\mathscr{C}^2_{(1,4)}$)

$$= baca[caa][ac][ca][ca]c(c)(c)(c)(ac)caacacccaac\varphi$$

$$= baca[caa][ac][ca][ca]c^2(c)(c)(ac)(ca)acacccaac\varphi$$

(using an instance of $\mathscr{C}^2_{(1,4)}$)

$$= baca[caa][ac][ca][ca]c(c)(c)(ac)(ca)acacccaac\varphi$$

$$= baca[caa][ac][ca][ca]c^2(c)(ac)(ca)(ac)acccaac\varphi$$

(using an instance of $\mathscr{C}^2_{(1,4)}$)

$$= baca[caa][ac][ca][ca]c(c)(ac)(ca)(ac)acccaac\varphi$$

$$= baca[caa][ac][ca][ca]c^2(ac)(ca)(ac)(ac)ccaac\varphi$$

(using an instance of $\mathscr{C}^2_{(1,4)}$)

$$= baca[caa][ac][ca][ca]c(ac)(ca)(ac)(ac)ccaac\varphi$$

$$= baca[caa][ac][ca][ca]ca(cca)(ac)(ac)(cca)ac\varphi$$

(using an instance of $\mathscr{D}^1_{(1,4)}$)

$$= baca[caa][ac][ca][ca]ac(cca)(ac)(ac)(cca)ac\varphi$$

$$= ba[ca][caa][acc][ac]a^2(ccca)(ac)(ac)(cca)ac\varphi$$

(using an instance of $\mathscr{C}^2_{(1,4)}$)

$$= ba[ca][caa][acc][ac]a(ccca)(ac)(ac)(cca)ac\varphi$$

$$= baca[caa][ac][ca][ca]c^2(c)(aac)(ac)(cca)ac\varphi$$

(using an instance of $\mathscr{C}^2_{(1,4)}$)

$$= baca[caa][ac][ca][ca]c(c)(aac)(ac)(cca)ac\varphi$$

$$= baca[caa][ac][ca][ca]c^2(aac)(ac)(cca)(ac)\varphi$$

(using an instance of $\mathscr{C}^2_{(1,4)}$)

$$= baca[caa][ac][ca][ca]c(aac)(ac)(cca)(ac)\varphi$$

$$= ba[ca][caa][ac][ca]ca(ca)(ac)(ac)(cca)ac\varphi$$

(using an instance of $\mathscr{D}^1_{(1,4)}$)

$$= ba[ca][caa][ac][ca]ac(ca)(ac)(ac)(cca)ac\varphi$$

$$= ba[ca][caa][ac][caa]c^2(aac)(ac)(cca)(ac)\varphi$$

(using an instance of $\mathscr{C}^2_{(1,4)}$)

$$= ba[ca][caa][ac][caa]c(aac)(ac)(cca)(ac)\varphi$$

$$= w''_3\varphi$$

(the segments inside [ ] are of type $l_2(9)$ in $\{a, c\}$ and the segments inside ( ) are either of type $r_1(1)$ in $\{c\}$ or of type $r_2(9)$ in $\{a, c\}$). Similarly, if we have some m'_i greater than $2m+1$. Hence $\mathscr{C}^{r+1}_{(1,m)}$ gives $w_m \dots w_1 a w'_1 \dots w'_m \varphi = w_m \dots w_1 a^2 w'_1 \dots w'_m \varphi$ and $f\varphi = g\varphi$ follows.

Now, if $u'' = d_1 \dots d_n$ (here, $uv = u'ad_1 \dots d_n v$ and $uav = u'ad_1 \dots d_n av$), then

$$a_1 f_1 a_2 f_2 \dots f_{i-1} a_i u' a d_1 \dots d_n v a_{i+1} f_{i+1} \dots a_{t-1} f_{t-1} a_t \varphi$$

$$= a_1 f_1 a_2 f_2 \dots f_{i-1} a_i u' a^2 d_1 \dots d_n v a_{i+1} f_{i+1} \dots a_{t-1} f_{t-1} a_t \varphi$$

(using in particular an instance of $\mathscr{C}^{r+1}_{(1,m)}$)

$$= a_1 f_1 a_2 f_2 \dots f_{i-1} a_i u' a d_1 a d_2 \dots d_n v a_{i+1} f_{i+1} \dots a_{t-1} f_{t-1} a_t \varphi$$

(using in particular an instance of $\mathscr{D}^r_{(1,m)}$)

$$= a_1 f_1 a_2 f_2 \dots f_{i-1} a_i u' a d_1 d_2 a d_3 \dots d_n v a_{i+1} f_{i+1} \dots a_{t-1} f_{t-1} a_t \varphi$$

(using in particular an instance of $\mathscr{D}^r_{(1,m)}$)

$$\vdots$$

$$= a_1 f_1 a_2 f_2 \dots f_{i-1} a_i u' a d_1 d_2 \dots d_{n-1} a d_n v a_{i+1} f_{i+1} \dots a_{t-1} f_{t-1} a_t \varphi$$

(using in particular an instance of $\mathscr{D}^r_{(1,m)}$)

$$= a_1 f_1 a_2 f_2 \dots f_{i-1} a_i u' a d_1 \dots d_n a v a_{i+1} f_{i+1} \dots a_{t-1} f_{t-1} a_t \varphi$$

(using in particular an instance of $\mathscr{D}^r_{(1,m)}$)

and $f\varphi = g\varphi$ follows. Let us show, for example, the equality

$$a_1 f_1 a_2 f_2 \dots f_{i-1} a_i u' a d_1 a d_2 d_3 \dots d_n v a_{i+1} f_{i+1} \dots a_{t-1} f_{t-1} a_t \varphi$$

$$= a_1 f_1 a_2 f_2 \dots f_{i-1} a_i u' a d_1 d_2 a d_3 \dots d_n v a_{i+1} f_{i+1} \dots a_{t-1} f_{t-1} a_t \varphi.$$

Using Lemma 2.1, rewrite $a_1 f_1 a_2 f_2 \dots f_{i-1} a_i u' a d_1$ as $w_m \dots w_1$ where

$$w_1, \dots, w_m \in A^* \text{ and } \{a, d_2\} \subseteq w_1\alpha \subseteq \cdots \subseteq w_m\alpha.$$

Also, rewrite $d_3 \dots d_n v a_{i+1} f_{i+1} \dots a_{t-1} f_{t-1} a_t$ as $w'_1 \dots w'_m$ where $w'_1, \dots, w'_m \in A^*$ and $\{a, d_2\} \subseteq w'_1\alpha \subseteq \cdots \subseteq w'_m\alpha$. We can choose w_1 to be the smallest suffix of $a_1 f_1 a_2 f_2 \dots f_{i-1} a_i u' a d_1$ to contain $\{a, d_2\}$, w_2 the smallest suffix of $a_1 f_1 a_2 f_2 \dots f_{i-1} a_i u' a d_1 - w_1$ to contain $w_1\alpha, \dots$ Similarly, we can choose w'_1 to be the smallest prefix of $d_3 \dots d_n v a_{i+1} f_{i+1} \dots a_{t-1} f_{t-1} a_t$ to contain $\{a, d_2\}$, w'_2 the smallest prefix of $d_3 \dots d_n v a_{i+1} f_{i+1} \dots a_{t-1} f_{t-1} a_t - w'_1$ to contain $w'_1\alpha, \dots$ Hence $w_m \dots w_1$ is in fact a suffix of $a_1 f_1 a_2 f_2 \dots f_{i-1} a_i u' a d_1$ and $w'_1 \dots w'_m$ is in fact a prefix of $d_3 \dots d_n v a_{i+1} f_{i+1} \dots a_{t-1} f_{t-1} a_t$. There exist nonnegative integers $n_0, \dots, n_r$ such that the n_0 first segments among $w_1, \ldots, w_m$ are of type $l_2(m_0)$ in $\{a, d_2\}$, the n_1 next segments are of type $l_3(m_1)$ in $\{a, d_2, b_1\}, \ldots$, and the last n_r segments are of type $l_{r+2}(m_r)$ in $\{a, d_2, b_1, \ldots, b_r\}$. Here, $m_0, \ldots, m_r$ are positive integers, $a, d_2, b_1, \ldots, b_r$ are in A. Similarly, there exist nonnegative integers $n'_0, \ldots, n'_r$ such that the n'_0 first segments among $w'_1, \ldots, w'_m$ are of type $r_2(m'_0)$ in $\{a, d_2\}$, the n'_1 next segments are of type $r_3(m'_1)$ in $\{a, d_2, c_1\}, \ldots$, and the last n'_r segments are of type $r_{r+2}(m'_r)$ in $\{a, d_2, c_1, \ldots, c_r\}$. Here, $m'_0, \ldots, m'_r$ are positive integers and $c_1, \ldots, c_r$ are in A. It is possible that some of $m_0, \ldots, m_r, m'_0, \ldots, m'_r$ be greater than $2m+1$. If this is the case for some m_i, say, $0 \leqslant i \leqslant r$, and w_k ($1 \leqslant k \leqslant m$) is of type $l_{i+2}(m_i)$ in an $(i+2)$-subset of A, one can write $w_k \varphi$ as $w''_k \varphi$ where w''_k is of type $l_{i+2}(2m+1)$ in that subset of A. This can be done using instances of $\mathscr{C}^{r+1}_{(1,m)}$ and $\mathscr{D}^{r}_{(1,m)}$. The result follows. $\square$

Now, let us reduce $\mathscr{C}^{r}_{(1,m)}$ and $\mathscr{D}^{r}_{(1,m)}$ to simpler equivalent sequences of equations.

Definition 4.2. Let $m \geqslant 1$ and let r be a nonnegative integer, $\mathscr{N}\mathscr{C}^{r}_{(1,m)}$ is the subsequence of $\mathscr{C}^{r}_{(1,m)}$ of all equations of the form

$$u_r \ldots u_0 x v_0 \ldots v_r = u_r \ldots u_0 x^2 v_0 \ldots v_r$$

where $u_0 = x^{n_0}$, $v_0 = x^{n'_0}$, where u_1 is the concatenation of n_1 segments of type $nl_2(2m+1)$ in $\{x, y_1\}$ (the last of which is of the form $x(y_1)^e$ for some $1 \leqslant e \leqslant 2m+1$), v_1 is the concatenation of n'_1 segments of type $nr_2(2m+1)$ in $\{x, z_1\}$ (the first of which is of the form $(z_1)^e x$ for some $1 \leqslant e \leqslant 2m+1$), where for $2 \leqslant i \leqslant r$, u_i is the concatenation of n_i segments of type $nl_{i+1}(2m+1)$ in $\{x, y_1, \ldots, y_i\}$ (the last of which starts with x if $n_1 = \cdots = n_{i-1} = 0$; otherwise, the last of which starts with a letter in $\{x, y_1, \ldots, y_j\}$ where j is the largest integer among $1, \ldots, i-1$ with $n_j \neq 0$), v_i is the concatenation of n'_i segments of type $nr_{i+1}(2m+1)$ in $\{x, z_1, \ldots, z_i\}$ (the first of which ends with x if $n'_1 = \cdots = n'_{i-1} = 0$; otherwise, the first of which ends with a letter in $\{x, z_1, \ldots, z_j\}$ where j is the largest integer among $1, \ldots, i-1$ with $n'_j \neq 0$), and where $n_i, n'_i \geqslant 0$, $0 \leqslant i \leqslant r$, and $m = n_0 + \cdots + n_r = n'_0 + \cdots + n'_r$. $\mathscr{N}\mathscr{D}^{r}_{(1,m)}$ is the subsequence of $\mathscr{D}^{r}_{(1,m)}$ of all equations of the form

$$u_r \ldots u_0 x y v_0 \ldots v_r = u_r \ldots u_0 y x v_0 \ldots v_r$$

where u_0 is the concatenation of n_0 segments of type $nl_2(2m+1)$ in $\{x, y\}$, v_0 is the concatenation of n'_0 segments of type $nr_2(2m+1)$ in $\{x, y\}$, where u_1 is the concatenation of n_1 segments of type $nl_3(2m+1)$ in $\{x, y, y_1\}$ (the last of which starts with a letter in $\{x, y\}$), v_1 is the concatenation of n'_1 segments of type $nr_3(2m+1)$ in $\{x, y, z_1\}$ (the first of which ends with a letter in $\{x, y\}$), where for $2 \leqslant i \leqslant r$, u_i is the concatenation of n_i segments of type $nl_{i+2}(2m+1)$ in $\{x, y, y_1, \ldots, y_i\}$ (the last of which starts with a letter in $\{x, y\}$ if $n_1 = \cdots = n_{i-1} = 0$; otherwise, the last of which starts with a letter in $\{x, y, y_1, \ldots, y_j\}$ where j is the largest integer among $1, \ldots, i-1$ with $n_j \neq 0$), v_i is the concatenation of n'_i segments of type $nr_{i+2}(2m+1)$ in $\{x, y, z_1, \ldots, z_i\}$ (the first of which ends with a letter in $\{x, y\}$ if $n'_1 = \cdots = n'_{i-1} = 0$; otherwise, the first of which ends with a letter in $\{x, y, z_1, \ldots, z_j\}$ where j is the largest integer among $1, \ldots, i-1$ with $n'_j \neq 0$), and where $n_i, n'_i \geqslant 0$, $0 \leqslant i \leqslant r$, and $m = n_0 + \cdots + n_r = n'_0 + \cdots + n'_r$.

Note that $\mathcal{N}\mathcal{C}^0_{(1,m)}$ consists of the equation $x^{2m+1} = x^{2m+2}$. We have $\mathcal{N}\mathcal{C}^0_{(1,m)} \subseteq \mathcal{N}\mathcal{C}^1_{(1,m)} \subseteq \mathcal{N}\mathcal{C}^2_{(1,m)} \subseteq \cdots$, and $\mathcal{N}\mathcal{D}^0_{(1,m)} \subseteq \mathcal{N}\mathcal{D}^1_{(1,m)} \subseteq \mathcal{N}\mathcal{D}^2_{(1,m)} \subseteq \cdots$.

Theorem 4.2. *$\mathcal{C}^r_{(1,m)}$ is equivalent to $\mathcal{N}\mathcal{C}^r_{(1,m)}$ and $\mathcal{D}^r_{(1,m)}$ is equivalent to $\mathcal{N}\mathcal{D}^r_{(1,m)}$. Consequently, $V_{2,1}$ is ultimately defined by $\bigcup_{r \geqslant 0} (\mathcal{N}\mathcal{C}^{r+1}_{(1,n)} \cup \mathcal{N}\mathcal{D}^r_{(1,n)})_{n \geqslant 1}$.*

Proof. The proof is similar to the one of Theorem 3.3. □

4.2. On a sequence of equations satisfied in $V_{2,n}$

In this section, for $n > 1$, we give a sequence of equations satisfied in $V_{2,n}$.

Definition 4.3. Let $n, m \geqslant 1$ and let r be a nonnegative integer. $\mathcal{C}^r_{(n,m)}$ is the finite sequence of all the equations in $\mathcal{C}^r_{(n)}$ where

$$\{x_1, \ldots, x_r\} = \{y_1, \ldots, y_r\} = \{z_1, \ldots, z_r\}$$

and where we replace

$$x \text{ by } (xx_1 \ldots x_r)^{\mathcal{N}} x (xx_1 \ldots x_r)^{\mathcal{N}},$$

$$y_i \text{ by } (xx_1 \ldots x_r)^{\mathcal{N}} y_i (xx_1 \ldots x_r)^{\mathcal{N}} \text{ and}$$

$$z_i \text{ by } (xx_1 \ldots x_r)^{\mathcal{N}} z_i (xx_1 \ldots x_r)^{\mathcal{N}}$$

for $1 \leqslant i \leqslant r$. Here $\mathcal{N} = (n+1)(m+1) - 1$ and is known to be the smallest N such that $x^N \sim_{(n,m)} x^{N+1}$ when $|x| = 1$.

We have $\mathcal{C}^0_{(n,m)} \subseteq \mathcal{C}^1_{(n,m)} \subseteq \mathcal{C}^2_{(n,m)} \subseteq \cdots$.

Theorem 4.3. *Let $|A| = r+1$, $r \geqslant 0$. Let M be a monoid generated by A. If M belongs to $V_{2,n}$, then M satisfies the equations in $\mathcal{C}^r_{(n,m)}$ for all m sufficiently large.*

Proof. It is easily seen, using Lemma 2.4, that $M \in V_{2,n}$ satisfies $\mathcal{C}^r_{(n,m)}$ for some $m \geqslant 1$. This comes from the fact that if $M \in V_{2,n}$, then M divides $A^*/\sim_{(n,m)}$ for some $m \geqslant 1$. Since $A^*/\sim_{(n,m)}$ satisfies $\mathcal{C}^r_{(n,m)}$, M satisfies $\mathcal{C}^r_{(n,m)}$. Moreover, M satisfies $\mathcal{C}^r_{(n,m')}$ for all $m' \geqslant m$ since $\sim_{(n,m')} \subseteq \sim_{(n,m)}$ for those m'. □

References

[1] F. Blanchet-Sadri, Some logical characterizations of the dot-depth hierarchy and applications, Ph.D. Thesis, McGill University, 1989.

[2] F. Blanchet-Sadri, Games, equations and the dot-depth hierarchy, *Comput. Math. Appl.* **18** (1989) 809–822.
[3] F. Blanchet-Sadri, On dot-depth two, *RAIRO Inform. Théor. Appl.* **24** (1990) 521–529.
[4] F. Blanchet-Sadri, Games, equations and dot-depth two monoids, *Discrete Appl. Math.* **39** (1992) 99–111.
[5] F. Blanchet-Sadri, Equations and dot-depth one, *Semi-group Forum*, to appear.
[6] F. Blanchet-Sadri, Some logical characterizations of the dot-depth hierarchy and applications, *Comput. System Sci.*, to appear.
[7] J.A. Brzozowski and R. Knast, The dot-depth hierarchy of star-free languages is infinite, *Comput. System Sci.* **16** (1978) 37–55.
[8] R.S. Cohen and J.A. Brzozowski, Dot-depth of star-free events, *Comput. System Sci.* **5** (1971) 1–15.
[9] S. Eilenberg, *Automata, Languages and Machines, Vol. B* (Academic Press, New York, 1976).
[10] S. Eilenberg and M.-P. Schützenberger, On pseudo-varieties, *Adv. in Math.* **19** (1976) 413–418.
[11] R. Knast, A semigroup characterization of dot-depth one languages, *RAIRO Inform. Théor.* **17** (1983) 321–330.
[12] R. Knast, Some theorems on graph congruences, *RAIRO Inform. Théor.* **17** (1983) 331–342.
[13] G. Lallement, *Semigroups and Combinatorial Applications* (Wiley, New York, 1979).
[14] D. Perrin and J.-E. Pin, First-order logic and star-free sets, *Comput. System Sci.* **32** (1986) 393–406.
[15] J.-E. Pin, *Variétés de Langages Formels* (Masson, Paris, 1984); *Varieties of Formal Languages* (Plenum, London, 1986).
[16] M.-P. Schützenberger, On finite monoids having only trivial subgroups, *Inform. and Control* **8** (1965) 190–194.
[17] I. Simon, Piecewise testable events, in: *Proc. 2nd GI Conf.*, Lecture Notes in Computer Science, Vol. 33 (Springer, Berlin, 1975) 214–222.
[18] H. Straubing, Finite semigroup varieties of the form $V*D$, *J. Pure Appl. Algebra* **36** (1985) 53–94.
[19] H. Straubing, Semigroups and languages of dot-depth two, in: *Proc. 13th ICALP*, Lecture Notes in Computer Science, Vol. 226 (Springer, New York, 1986) 416–423.
[20] W. Thomas, Classifying regular events in symbolic logic, *Comput. System Sci.* **25** (1982) 360–376.
[21] W. Thomas, An application of the Ehrenfeucht–Fraïssé game in formal language theory, *Bull. Soc. Math. de France* **16** (1984) 11–21.
[22] W. Thomas, A concatenation game and the dot-depth hierarchy, in: E. Börger, ed., Computation Theory and Logic, Lecture Notes in Computer Science, Vol. 270 (Springer, Berlin, 1987) 415–426.
[23] P. Weil, Inverse monoids of dot-depth two, *Theoret. Comput. Sci.* **66** (1989) 233–245.